

Können Algorithmen diskriminieren?

Wiebke Fröhlich

2018-12-26T10:02:40

In Österreich kategorisiert ab 2019 ein Algorithmus arbeitslose Personen nach ihren Chancen auf dem Arbeitsmarkt. Die [Software trennt in drei Personengruppen](#): Arbeitssuchende mit guten, mittleren und schlechten Perspektiven, einen Arbeitsplatz zu finden. Auf dieser Basis will der [Arbeitsmarktservice Österreich \(AMS\)](#) seine Ressourcen ab 2020 überwiegend auf Personen der mittleren Gruppe konzentrieren. Dort seien sie am effektivsten eingesetzt. Die „Arbeitsmarktintegrationschancen“ von [Frauen bewertet der Algorithmus pauschal negativ](#). Zudem führen betreuungspflichtige Kinder zu einer schlechten Einstufung – allerdings nur für Frauen. Bei Männern, so [begründen die Entwickler](#), habe eine Betreuungspflicht statistisch gesehen keine negativen Auswirkungen auf die Arbeitsmarktchancen.

Festschreibung bestehender Missstände

Algorithmen wie der hier beschriebene generieren aus großen Datenmengen – im Idealfall statistisches – Wissen, um Informationslücken zu schließen. Aus vergangenheitsbezogenen Informationen werden Rückschlüsse auf Eigenschaften und künftiges Verhalten von Personen gezogen, deren eigene Daten der Algorithmus nicht zwingend ausgewertet hat. Anhand von Stellvertretermerkmalen (z.B. „Frau“) werden Personen Eigenschaften zugeschrieben, die Trägerinnen des Merkmals vermeintlich überwiegend und typischerweise aufweisen (z.B. schlechte Arbeitsmarktintegrationschancen).

Ist der Algorithmus des AMS ein Paradebeispiel eines „[diskriminierenden Algorithmus](#)“? Die Befürworterinnen entsprechender Techniken preisen Effizienz und Genauigkeit bei Ressourcenverteilungs- und Prognoseentscheidungen. Der Computer entscheide objektiv, jedenfalls weniger voreingenommen als Menschen. Wenn der Algorithmus „Frau-Sein“ negativ bewerte, stelle er nur die strukturell bedingten Nachteile von Frauen im Erwerbsleben dar, [argumentiert beispielsweise der AMS-Behördenleiter](#). Ignoriere man die Tatsache, dass bestimmte Gruppen auf dem Arbeitsmarkt diskriminiert würden, könne man diese auch nicht korrekt fördern. Dahinter steht eine Ressourcenentscheidung: Da nicht alle gefördert werden können, werden diejenigen privilegiert, bei denen der Algorithmus das Geld am effektivsten investiert sieht.

Kritikerinnen lassen das nicht gelten: Wenn die soziale Realität diskriminierend sei, wirken [Algorithmen, die diese Realität abbilden, ebenfalls diskriminierend](#). Durch die Wiedergabe der Realität (Benachteiligung von Frauen am Arbeitsmarkt; Betreuungspflichten – bei Frauen – als Beschäftigungsnachteil) wird ein [bestehender Missstand festgeschrieben](#). Im Zweifel wird das Übel sogar verstärkt. Denn wenn eine bereits benachteiligte Person aufgrund der Einschätzung des Algorithmus nicht gefördert wird, verschlechtern sich ihre Chancen weiterhin.

Zum Vorschein kommt ein Dilemma: Wer Diskriminierung bekämpfen will, muss diese sichtbar machen (oder zumindest selbst sehen). Gleichzeitig kann die statistische Abbildung diskriminierender Strukturen zu einer selbsterfüllenden Prophezeiung werden und Diskriminierungen verstärken. Nicht der Algorithmus diskriminiert, sondern die Stelle, die ihn einsetzt.

Die Insuffizienz von Diskriminierungsverboten

Zwar setzt das hiesige [Jobcenter laut Bundesagentur für Arbeit](#) (noch) keine Algorithmen ein. Doch auch in Deutschland herrscht in Ämtern und Behörden Personalmangel und Ressourcenknappheit. Die Vorstellung, mit computergestützten Verfahren Prozesse beschleunigen und die Effizienz zu steigern, wird in [Politik](#) (z.B. [Datenethikkommission der Bundesregierung](#) sowie [Enquete-Kommission „Künstliche Intelligenz“](#)) und Wissenschaft (z.B. nur [hier](#) und [hier](#)) diskutiert. Positiv ist, dass Diskriminierung als Problem algorithmengestützter Entscheidungen erkannt ist. Als Lösung werden insbesondere [antidiskriminierungsrechtlichen Maßnahmen diskutiert](#). Dass Diskriminierungsverbote aber nur bedingt als Instrumente einer Algorithmenregulierung taugen, wird am österreichischen Beispiel besonders deutlich. Diskriminierungsverbote setzen eine „Benachteiligung“ bzw. „benachteiligende Behandlung“ voraus. Justiziabel ist erst die auf einer algorithmischen Datenanalyse basierende *Entscheidung*. Den algorithmischen Befund als solchen tasten Diskriminierungsverbote nicht an --- auch wenn (vor allem) auf ihm die Entscheidung gründet.

Jedoch steht der AMS-Algorithmus nicht in der Kritik, weil er etwa dazu führt, dass Frauen weniger gefördert werden als Männer. Abzulehnen ist schon die negative Bewertung des „Frau-Seins“ als solche, nämlich die algorithmische Kategorisierung von Personen im Vorfeld der Förderentscheidung anhand dieser Kriterien. Diese dient dazu, Menschen anhand einer statistischen Norm zu bewerten und ihnen darauf basierend Eigenschaften zuzuschreiben. Diese Betrachtungsweise ist nur scheinbar objektiv. Tatsächlich unterliegen ihr normative Annahmen – insbesondere die, dass das Geschlecht ein solides Differenzierungskriterium ist.

Externe Persönlichkeitskonstruktion durch Zuschreibung von Gruppenmerkmalen als rechtliches Problem

Denn es wird regelmäßig übersehen, dass bereits die Zuschreibung von Eigenschaften im Vorfeld einer Entscheidung grundrechtliche Relevanz hat. Die externe Persönlichkeitskonstruktion als solche (also die Charakterisierung anhand gruppenbildender Merkmale durch Algorithmen) greift in das Recht auf informationelle Selbstbestimmung (nach Art. 1 Abs. 1, 2 Abs. 1 GG) bzw. das ähnlich motivierte Recht auf Datenschutz (Art. 7 bzw. Art. 8 EU-GRCh) ein. Dieses gewährleistet als Ausprägung des Rechts auf Selbstdarstellung „die Freiheit des Einzelnen, selbst zu bestimmen, welches Persönlichkeitsbild er von sich vermitteln

will“ (vgl. für das deutsche Recht [BVerfGE 82, 236 \(269\)](#)). Darüber hinaus dürften auch andere Grundrechte betroffen sein, um die es hier nicht weiter gehen soll.

Computerprogramme, die statistische Persönlichkeitsbilder erstellen, verkürzen die Freiheit auf informationelle Selbstdarstellung: Wer von einem Algorithmus in eine statistische Schublade gesteckt wird, kommt aus dieser kaum mehr heraus. Eine Sachbearbeiterin, welcher der AMS-Algorithmus sagt, dass Frauen generell schlechte Chancen auf dem Arbeitsmarkt haben, wird jeder Bewerberin mit dieser Voreinstellung begegnen und ihr nur ausnahmsweise Förderung gewähren. Es entsteht ein (negatives) Bild der Person, dem diese nichts entgegensetzen kann. Dieses Persönlichkeitsbild wird nicht in einem kommunikativen und grundsätzlich gleichgewichtigen Prozess unter Möglichkeit der Einflussnahme der Betroffenen gewonnen, sondern von Dritten anhand Daten Dritter konstruiert.

Dass Persönlichkeitskonstruktionen durch *automatische* Datenverarbeitung besonders problematisch sind, hat das BVerfG bereits 1983 erkannt. In seinem [Volkszählungsurteil](#) heißt es:

„Das [...] allgemeine Persönlichkeitsrecht, [...] umfasst [...] auch die aus dem Gedanken der Selbstbestimmung folgende Befugnis des Einzelnen, grundsätzlich selbst zu entscheiden, wann und innerhalb welcher Grenzen persönliche Lebenssachverhalte offenbart werden [...]. Diese Befugnis bedarf unter den heutigen und künftigen Bedingungen der automatischen Datenverarbeitung in besonderem Maße des Schutzes.“ (BVerfGE 65, 1 (41f.))

Die damals ausgemachte Bedrohungslage hat durch die Entwicklung der Informationstechnologie eine noch viel weiterreichende Dimension erhalten. 1983 ging die Gefahr von der Verarbeitung und Rekombination vieler Einzelangaben einer konkreten Person aus. [Weder der Staat noch Private](#) begnügen sich heute aber mit den Daten einer Person, um deren Persönlichkeitsbild zu konstruieren. Vielmehr ermitteln Algorithmen Gruppeninformationen und schreiben sie der Einzelnen zu. Ein solches Persönlichkeitsbild kann das Recht auf Selbstdarstellung mindestens ebenso gewichtig beeinträchtigen wie die Auswertung vieler Einzelangaben einer Person. Zum einen ignoriert die Zuweisung von Gruppeneigenschaften die Individualität der Gruppenmitglieder. Zum anderen wird den Personen die Möglichkeit genommen, selbst über „Ob“ und „Wie“ der Preisgabe der Eigenschaft zu entscheiden – oder den Befund des Gegenübers zu beeinflussen. Das gilt insbesondere, wenn das Rechenergebnis an Merkmale anknüpft, die bereits mit Stigmata und Stereotypen behaftet sind.

Der [AMS-Algorithmus unterstellt](#), dass „Frauen“ oder „Frauen mit Betreuungspflicht“ eine homogene Bevölkerungsgruppe sind. Indem der Algorithmus nach Geschlecht differenziert, wird das eigentliche Problem verfestigt. Denn eine daran orientierte, staatliche Ressourcenverteilung legitimiert das Geschlecht als Differenzierungsmerkmal. Dies widerspricht sowohl dem verfassungsrechtlichen Ziel (Art. 3 Abs. 2 S. 2 GG) der tatsächlichen Gleichstellung von Mann und Frau (vgl. auch Art. 23 EU-GRCh) als auch der durch Art. 2 Abs. 1 S. 1 GG bzw. Art. 7, 8 EU-GRCh geschützten freien Entfaltung der individuellen Persönlichkeit.

Datenschutzrecht und Regulierung informationeller und informationstechnischer Macht

Eine Algorithmenregulierung darf daher nicht erst an – gegebenenfalls – diskriminierenden Verteilungsentscheidungen ansetzen, sondern muss bereits auf der Vorfeldebene intervenieren. Ein probates Mittel zur Durchbrechung algorithmischer Persönlichkeitskonstruktion ist das Datenschutzrecht, das Personen in ihrem Recht auf informationelle Selbstbestimmung vor informationellen Machtasymmetrien schützt. Das BVerfG hat das Recht auch und gerade zum Schutz der freien Persönlichkeitsentfaltung in kommunikativen Prozessen und der darauf aufbauenden gesamtgesellschaftlichen Freiheitlichkeit konzipiert:

„Wer unsicher ist, ob abweichende Verhaltensweisen [...] gespeichert, verwendet oder weitergegeben werden, wird versuchen, nicht durch solche Verhaltensweisen aufzufallen. [...] Dies würde nicht nur die [...] Entfaltungschancen des Einzelnen beeinträchtigen, sondern auch das Gemeinwohl, weil Selbstbestimmung eine elementare Funktionsbedingung eines auf Handlungsfähigkeit und Mitwirkungsfähigkeit seiner Bürger begründeten freiheitlichen demokratischen Gemeinwesens ist.“ ([BVerfGE 65, 1 \(43\)](#))

Der [überindividuelle Wert von Entfaltungsfreiheit](#) ist längst anerkannt. Auch die [Anfänge der einfachen Datenschutzgesetzgebung in den 1970er Jahren](#) gehen auf den Bedarf zurück, strukturelle informationelle Macht zu regulieren. Gleichwohl beschränkt sich das Datenschutzrecht bis heute, auch in der DSGVO, vor allem auf eine Kontrolle der Verarbeitung von Daten. Dabei waren die Bedeutung der Daten und der Technologie für Entscheidungen und ihre Folgen immer im Blick. Wenn Daten aber im Zeitalter von Big Data nicht mehr beherrschbar sind, muss eine Ergebniskontrolle in den Fokus rücken.

Verbot automatisierter Entscheidungen im Einzelfall

Die DSGVO kennt nur wenige Vorschriften, die sich mit Entscheidungen aufgrund von Datenverarbeitungen befassen, vor allem Art. 22 DSGVO. Dort heißt es in Abs. 1: „Die betroffene Person hat das Recht, nicht einer ausschließlich auf einer automatisierten Verarbeitung [...] beruhenden Entscheidung unterworfen zu werden [...]“. Diesen Grundsatz durchlöchern allerdings die weitreichenden Ausnahmen in Abs. 2. Zudem greift das Recht des ersten Absatzes nur, wenn die Entscheidung als solche automatisiert erfolgt. Den Arbeitssuchenden in Österreich hilft Art. 22 DSGVO also kaum. Denn die Entscheidung über die Zuteilung von Fördermaßnahmen trifft letztlich die Sachbearbeiterin, also ein Mensch; der Algorithmus unterstützt sie „nur“.

Es steht aber zu befürchten, dass Sachbearbeiterinnen die Empfehlungen des Computers ohne Weiteres übernehmen werden. Die eigentliche Gefährdung automatisiert unterstützter Entscheidungsfindung besteht in einer faktischen Umkehr der Argumentations- und womöglich sogar Beweislast, der das Individuum nicht gerecht werden kann: Die arbeitssuchende Frau muss aufzeigen, dass sie „anders“

ist als der algorithmisch ermittelte Wert. Dabei kann sie nicht wissen, welche Daten der Algorithmus wie bewertet hat. Und selbst wenn sie es wüsste, wäre es ihr faktisch nahezu unmöglich, dieses Ergebnis mit anderen Bewertungen und Grundlagen technisch und inhaltlich anzugreifen. Betroffenenrechte laufen faktisch leer.

Keine Neutralität von Algorithmen

Am Anfang einer Algorithmenregulierung muss die Erkenntnis stehen, dass Algorithmen nicht neutral sind. Schon indem sie an bestimmte Merkmale anknüpfen und Korrelationen zu anderen Merkmalen ermitteln, nehmen sie eine Wertung vor. Diese Wertung schlägt auf die Entscheidung durch. Das schließt einen Einsatz von Algorithmen nicht etwa aus. Zu Recht argumentieren die Befürworterinnen der Technik, dass – wie im Fall AMS – gerade durch das Sichtbarmachen statistisch belegbarer Unterschiede eine effektivere Verteilung staatlicher Ressourcen gelingen kann. Umgekehrt liegt aber gerade darin die Gefahr einer Perpetuierung bestehender Missstände und die Verwirklichung selbsterfüllender Prophezeiungen.

Es ist Aufgabe des Sozialstaats, der einzelnen Person in ihrer Individualität und Würde gerecht zu werden und wider statistische Erfahrungswerte zu handeln, die immer Werte der Mehrheit sind. Es ist nicht zu verhehlen, dass sowohl der Sozialstaat bei der Verteilung knapper Ressourcen als auch Arbeitgeberinnen/ Unternehmen, die sich durch Algorithmen in ihren Entscheidungen unterstützen lassen, gewichtige berechnete Interessen ins Feld führen können. Insbesondere bei Entscheidungen über elementare Grundbedürfnisse wie Arbeit, Sozialhilfe, Gesundheit sind aber die Gefahren diskriminierender Entscheidungen bereits auf Ebene der Datenverarbeitung mitzudenken.

Fehlen rechtliche Regelungen für den Einsatz von Algorithmen, können Beschränkungen darauf basierender Datenverarbeitung aus dem Antidiskriminierungsrecht abgeleitet werden. Art. 23 UAbs. 1 EU-GRCh verpflichtet die Mitgliedstaaten, die Gleichheit von Frauen und Männern sicherzustellen. Auch in Art. 7 Abs. 2 der österreichischen Bundesverfassung bekennt sich der Staat zur tatsächlichen Gleichstellung von Mann und Frau. In Deutschland fordert Art. 3 Abs. 2 S. 2 GG dazu auf, die tatsächliche Durchsetzung der Gleichberechtigung von Frauen und Männern zu fördern.

Weitergedacht könnten die Gleichstellungsgebote Verbote statistischer Datenauswertungen für besonders diskriminierungsanfällige Lebensbereiche legitimieren. Dem entspricht auch das schon bestehende [Verbot geschlechtsabhängiger Versicherungstarife](#). Um mittelbare Beeinträchtigungen, wie etwa die Unterscheidung nach Teil- und Vollzeitbeschäftigung (überproportional viele Teilzeitbeschäftigte sind Frauen) zu verhindern, könnten die Verarbeitungsverbote auch [auf Stellvertretermerkmale auszuweiten](#) sein. Zudem ist der Blick auf die Rechtsdurchsetzung zu lenken: Einzelne können algorithmische Gruppenaussagen kaum erschüttern. Individualisierter Rechtsschutz vermag das kollektive Risiko des Einsatzes von Algorithmen nicht hinreichend aufzufangen. Aus den materiellen Rechtsregeln könnten daher sogar verfahrensrechtliche Forderungen abzuleiten

sein. Es ist bei weitem nicht (nur) Aufgabe des Datenschutzrechts, sondern des Rechts insgesamt, die Diskriminierungspotentiale des Einsatzes von Algorithmen normativ einzuhegen und dies schon auf der Ebene der Entscheidungsvorbereitung. Nicht Algorithmen diskriminieren, sondern die, die sie ohne Hinterfragen entwickeln oder einsetzen.

